

Grand Challenge Project

D. Olson, LBNL

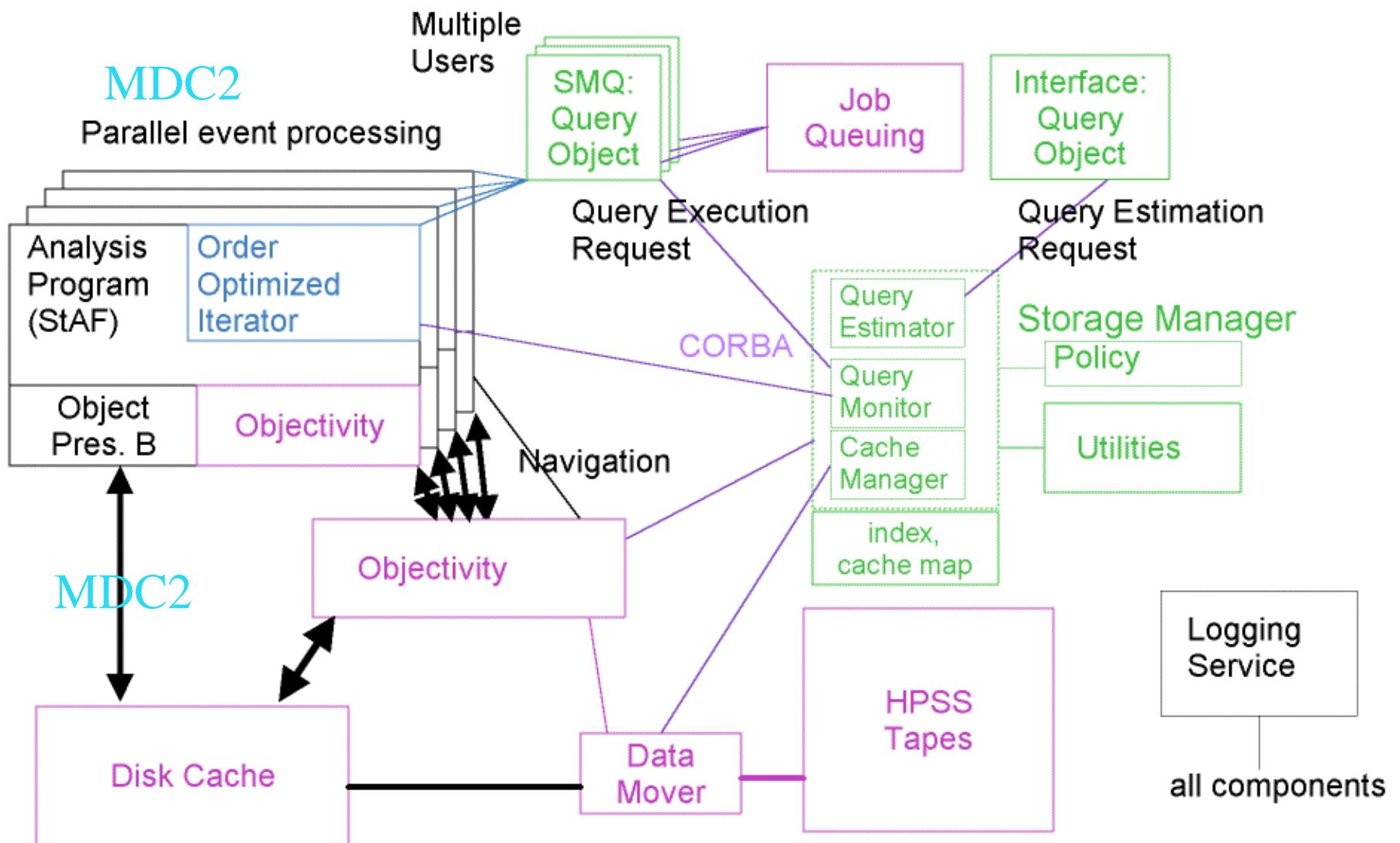
RHIC Computing Advisory Committee Meeting

4 Dec 1998, BNL

Outline

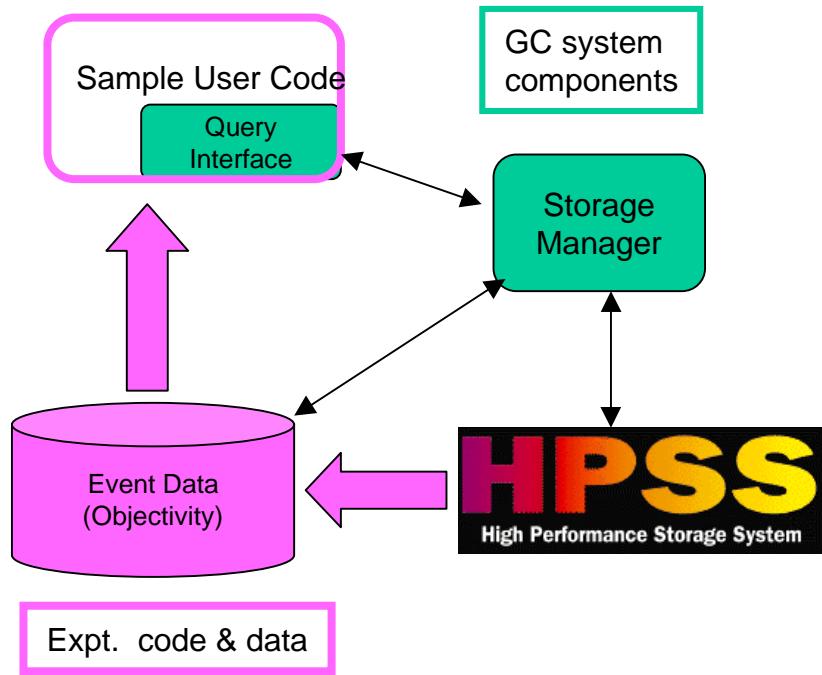
- Overview
- MDC1 features & results
- MDC2 plans
- Summary

RHIC Analysis Architecture



D. Olson, Dec 97

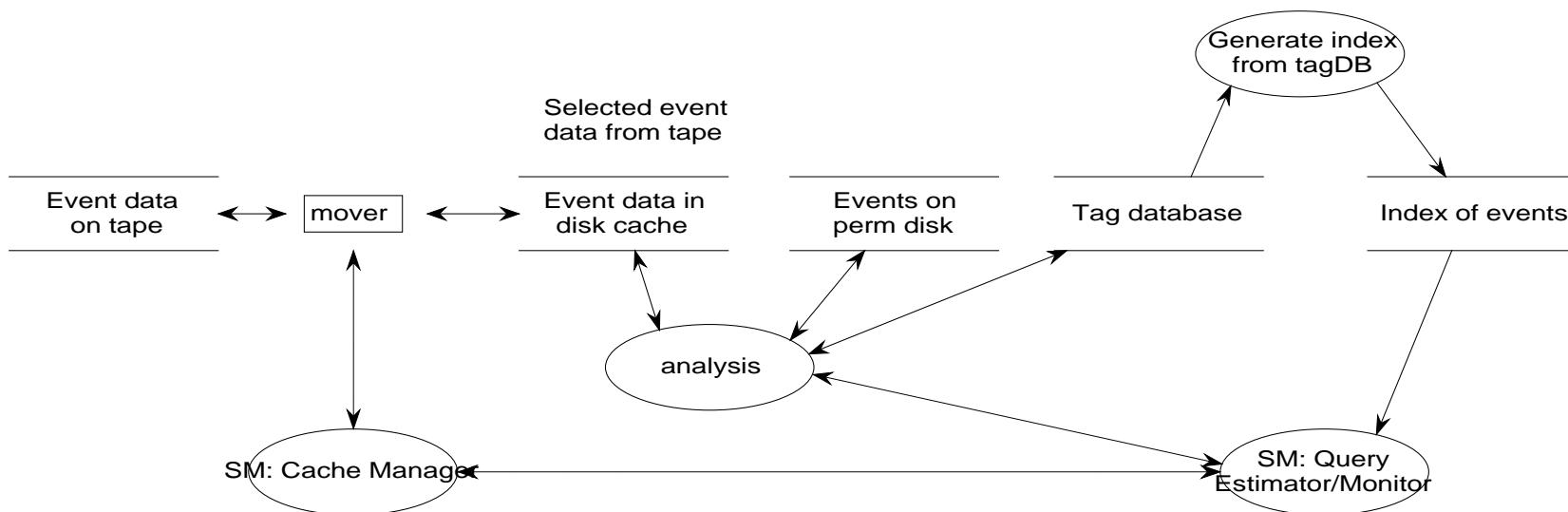
Grand Challenge software



- key developers

- Henrik Nordberg (NERSC)
query estimator
- Alex Sim (NERSC)
query monitor
- Luis Bernardo (NERSC)
cache manager
- Jeff Porter (LBL-STAR)
query object
- Dave Malon (ANL)
order-optimized iterator &
gcaResources API
- Dave Zimmerman, (LBL-STAR)
Mark Pollack (BNL-PHENIX)
tagDB
- Jie Yang (UCLA,LBL,Beijing)
testing

Data Flow View



OMT, RA1
29 Mar 98

MDC1

- Single client analysis process per query
- Multiple simultaneous queries with shared file policy
- Single event component per event ID
 - event is stored in one file, not split across files
- Use Objectivity/DB for event ID, tagDB, event data
- Scale of tests at <1% of nominal year

Sample tag data & queries

Tag_Flat_Index.data

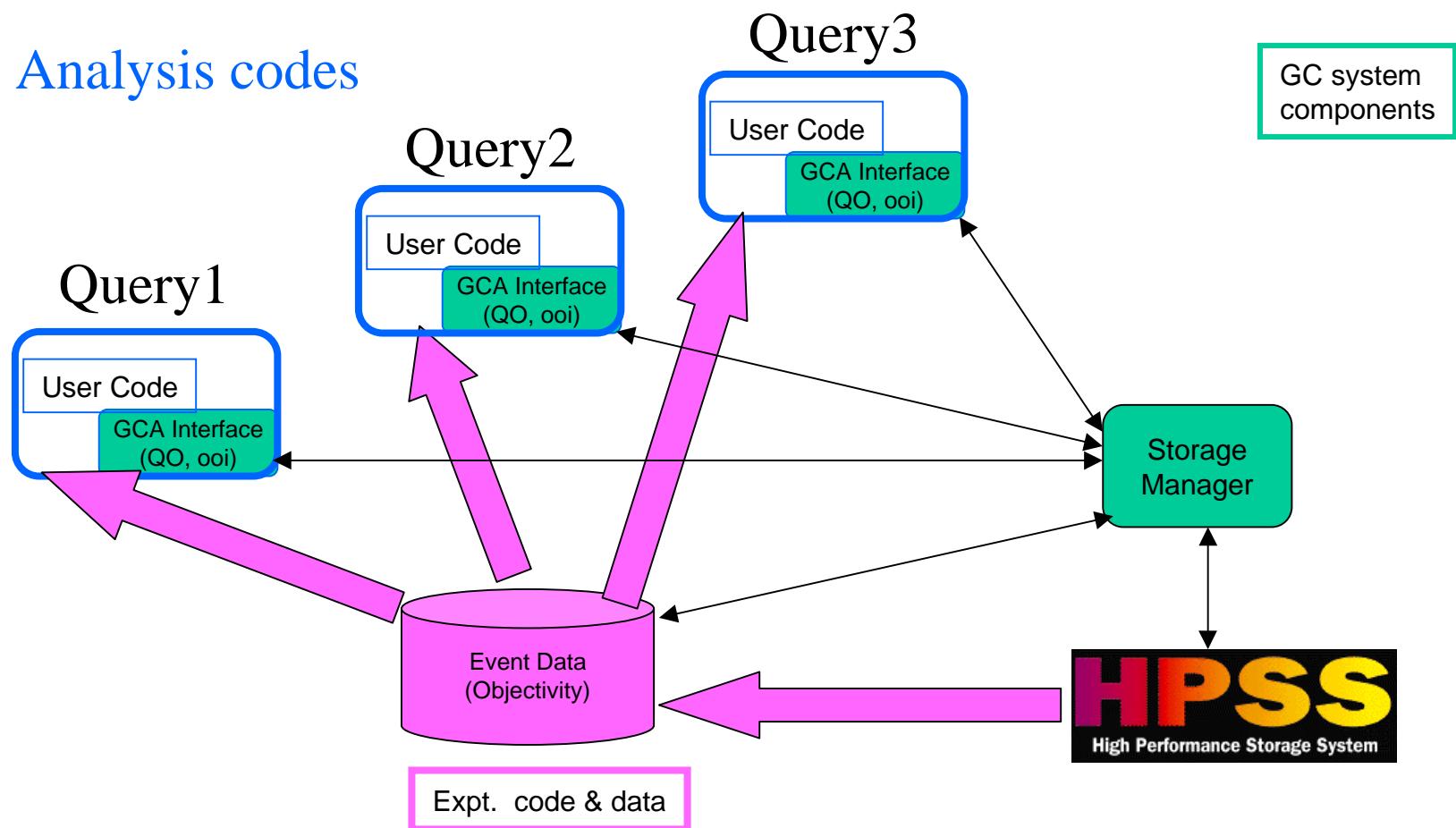
```
n_event0 n_event1 prod_run glb_trk_tot glb_trk_good glb_trk_prim glb_trk_plus g  
lb_trk_minus n_vert_total n_vert_V0 n_vert_K0 n_vert_Lambda n_vert_ALambda n_ver  
t_pileup mult_eta0 mult_eta1 mult_eta2 mult_eta3 mult_eta4 mult_pt0 mult_pt1 mul  
t_pt2 mult_pt3 mult_pt4 mult_phi0 mult_phi1 mult_phi2 mult_phi3 mult_phi4 mean_p  
t mean_eta T_average T_eta_bins0 T_eta_bins1 T_eta_bins2 energy_emc_eta0 energy_  
emc_eta1 energy_emc_eta2 energy_emc_eta3 energy_emc_eta4 energy_emc_phi0 energy_  
_emc_phi1 energy_emc_phi2 energy_emc_phi3 energy_emc_phi4 prim_vrtx0 prim_vrtx1  
prim_vrtx2 prim_vrtx_chisq tagrand nTags EventID.dbaseID EventID.contID EventID  
.pageID EventID.slotID Set file_ID  
0 0 0 5845 5388 0 3114 2731 50206 50205 50205 0 0 0 393 1283 1969 1349 394 3140  
1433 459 186 77 147 131 651 1094 1029 0.391864 0.012653 0.175159 0.179548 0.173  
942 0.179345 0 0 0 0 0 0 0 -0.00075826 0.00654174 0.00436799 0 665.67 52 3  
2 3 5 77 7  
0 0 0 6201 5702 0 3279 2922 54783 54782 54782 0 0 0 433 1311 2190 1362 406 3312  
1486 508 184 83 151 125 650 1114 1099 0.396015 -0.00380799 0.178561 0.192104 0.  
169312 0.180685 0 0 0 0 0 0 0 0 -0.00260277 -0.00826739 -0.00690387 0 620.04  
5 52 3 2 3 31 77 7
```

Tag attribute
names
values

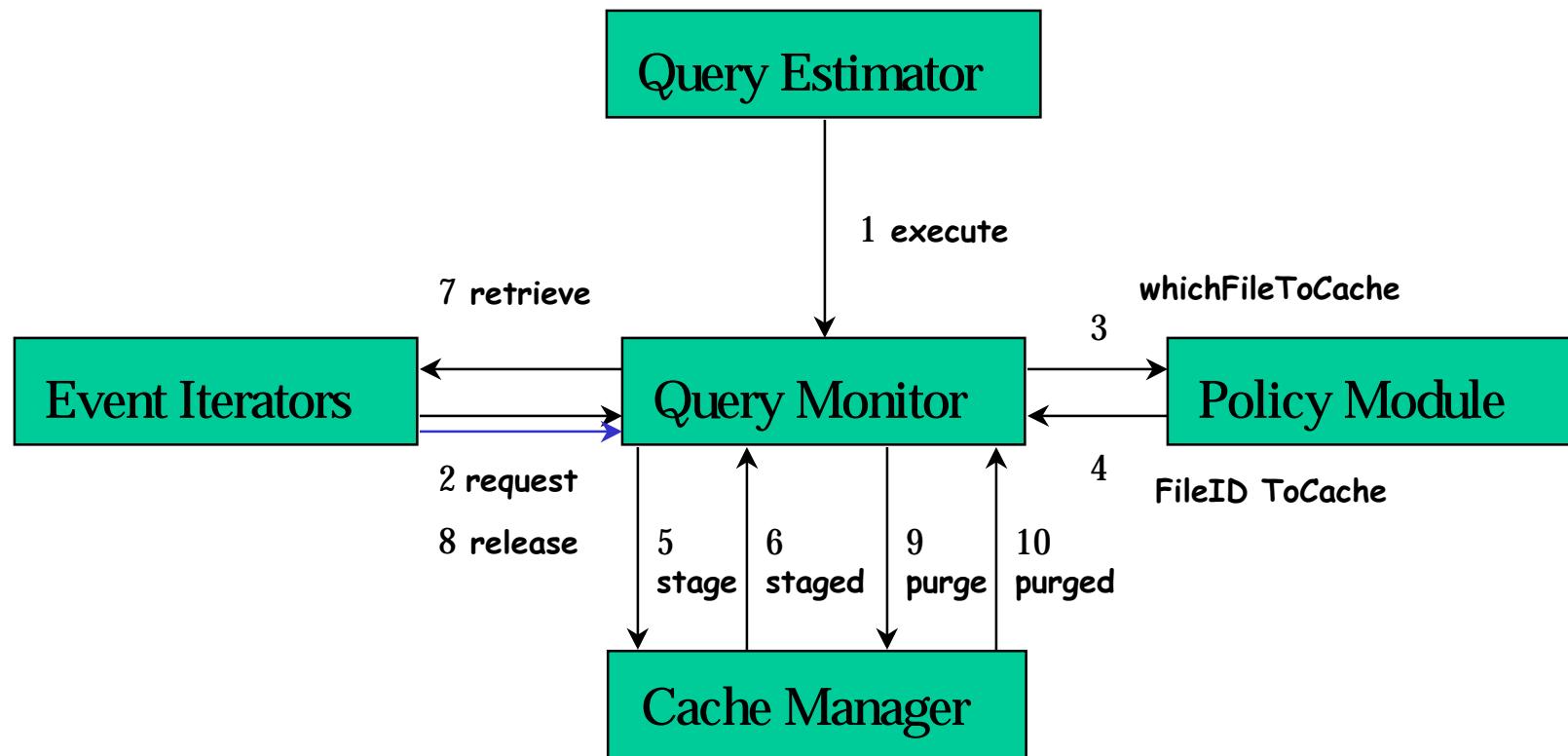
Sample queries

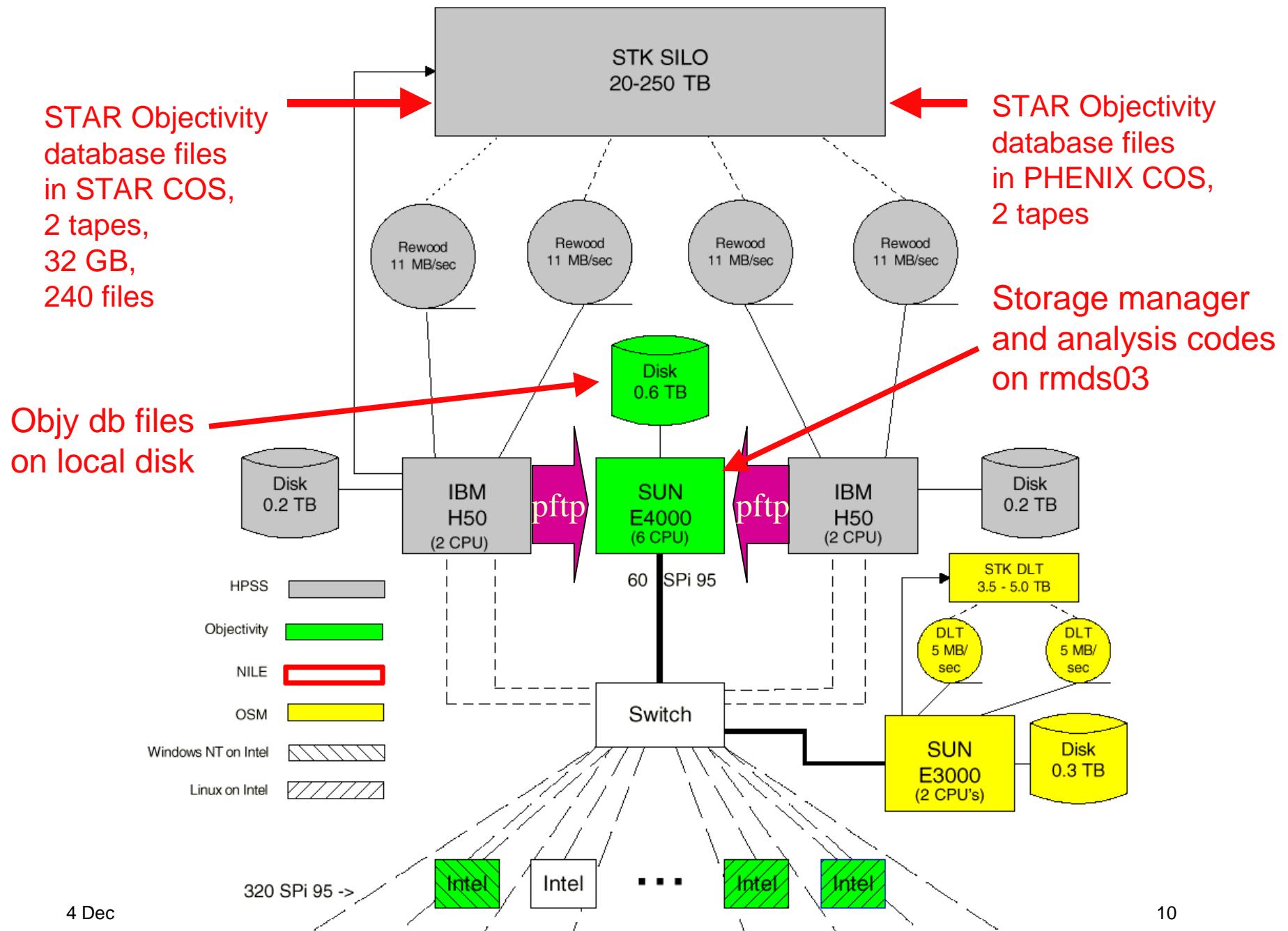
```
"mean_eta<-0.2" : 0.1 : 5  
"glb_trk_tot>5000&mean_pt>0.5&tagrand>1000." : 0.1 : 5  
"glb_trk_tot>0&glb_trk_tot<10&n_vert_total<3" : 0.1 : 5  
"glb_trk_tot>0&glb_trk_tot<10&mean_eta>0.2" : 0.1 : 10000
```

Multiple simultaneous queries

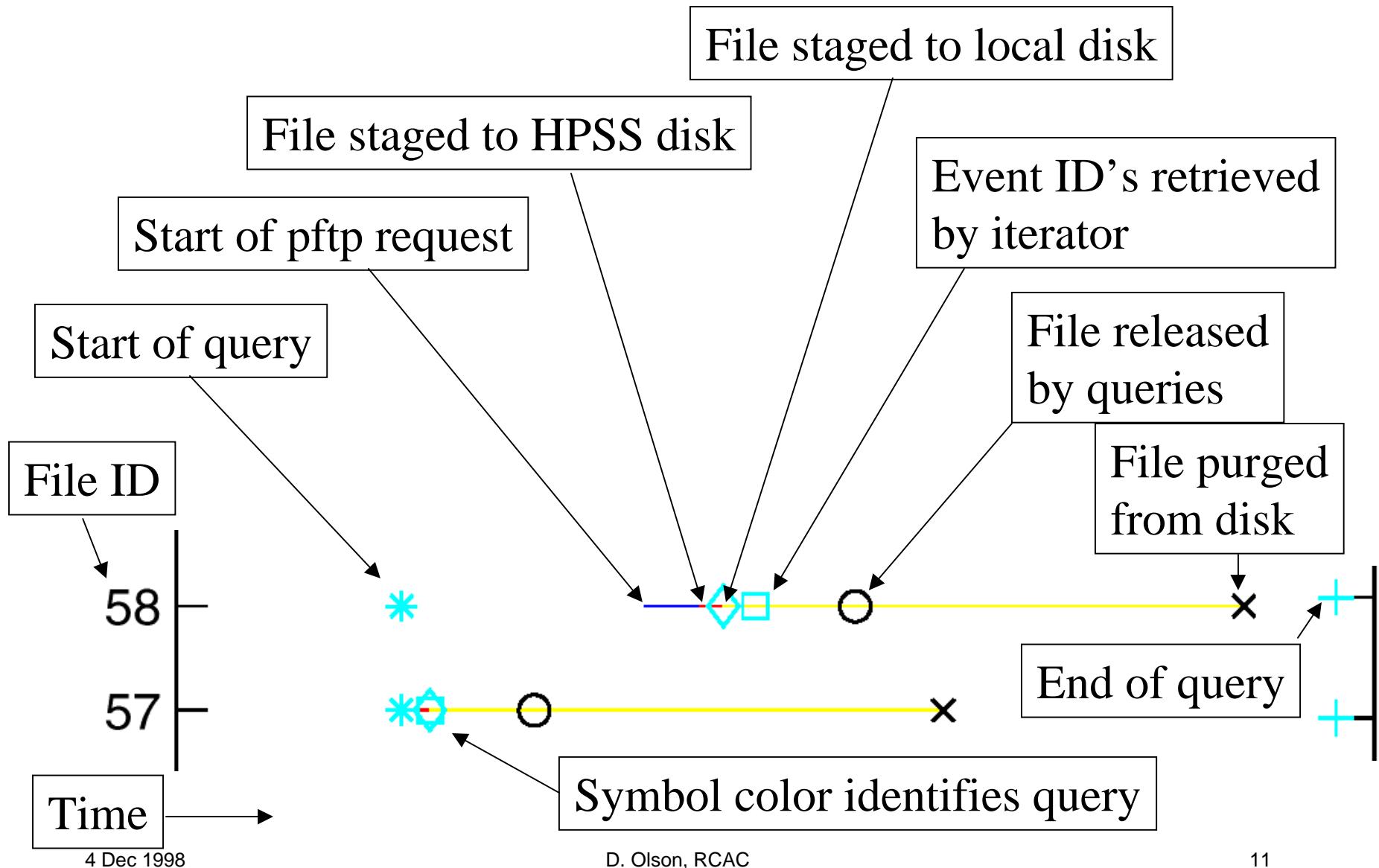


MDC1 Process Flow





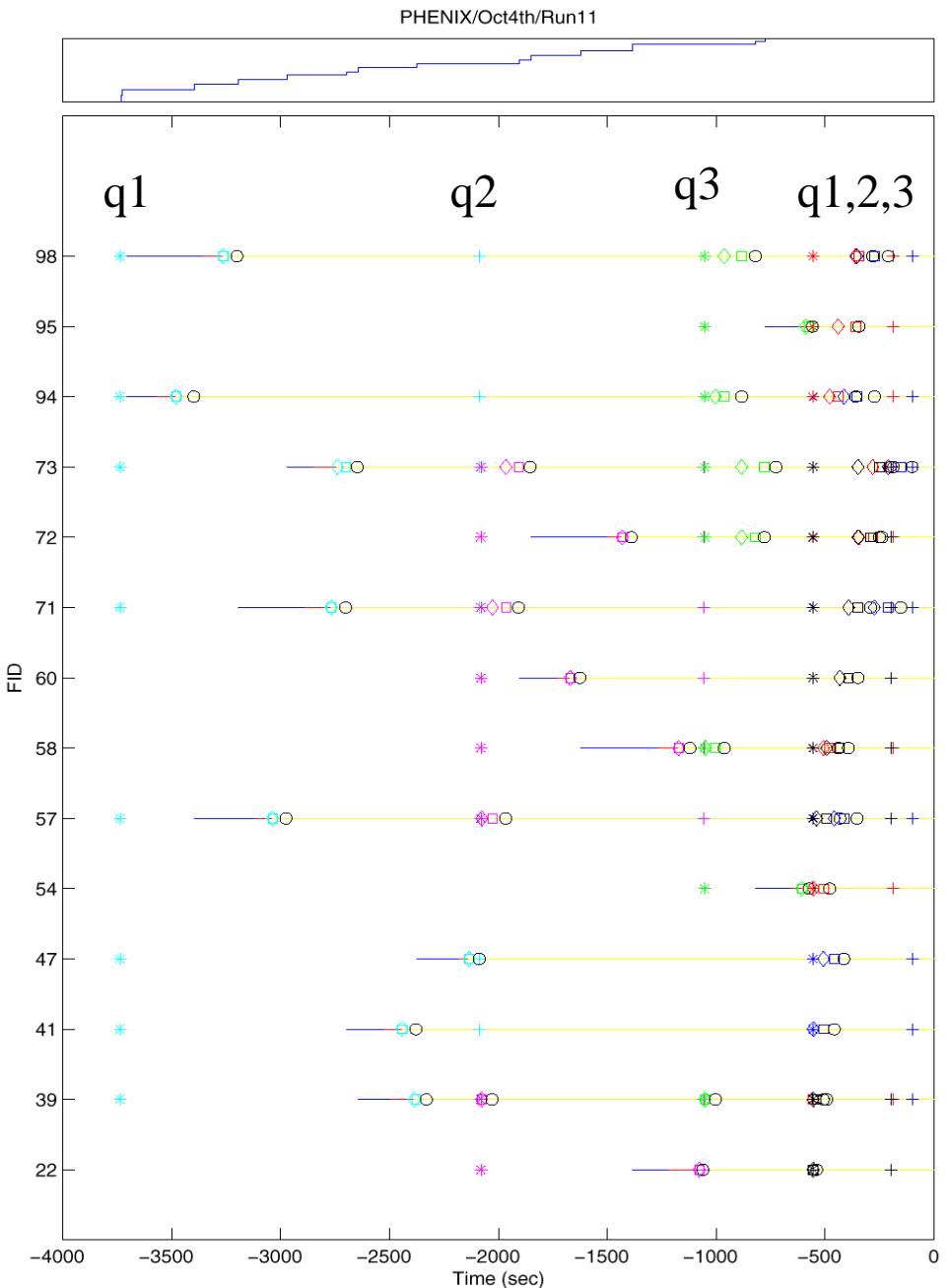
Legend



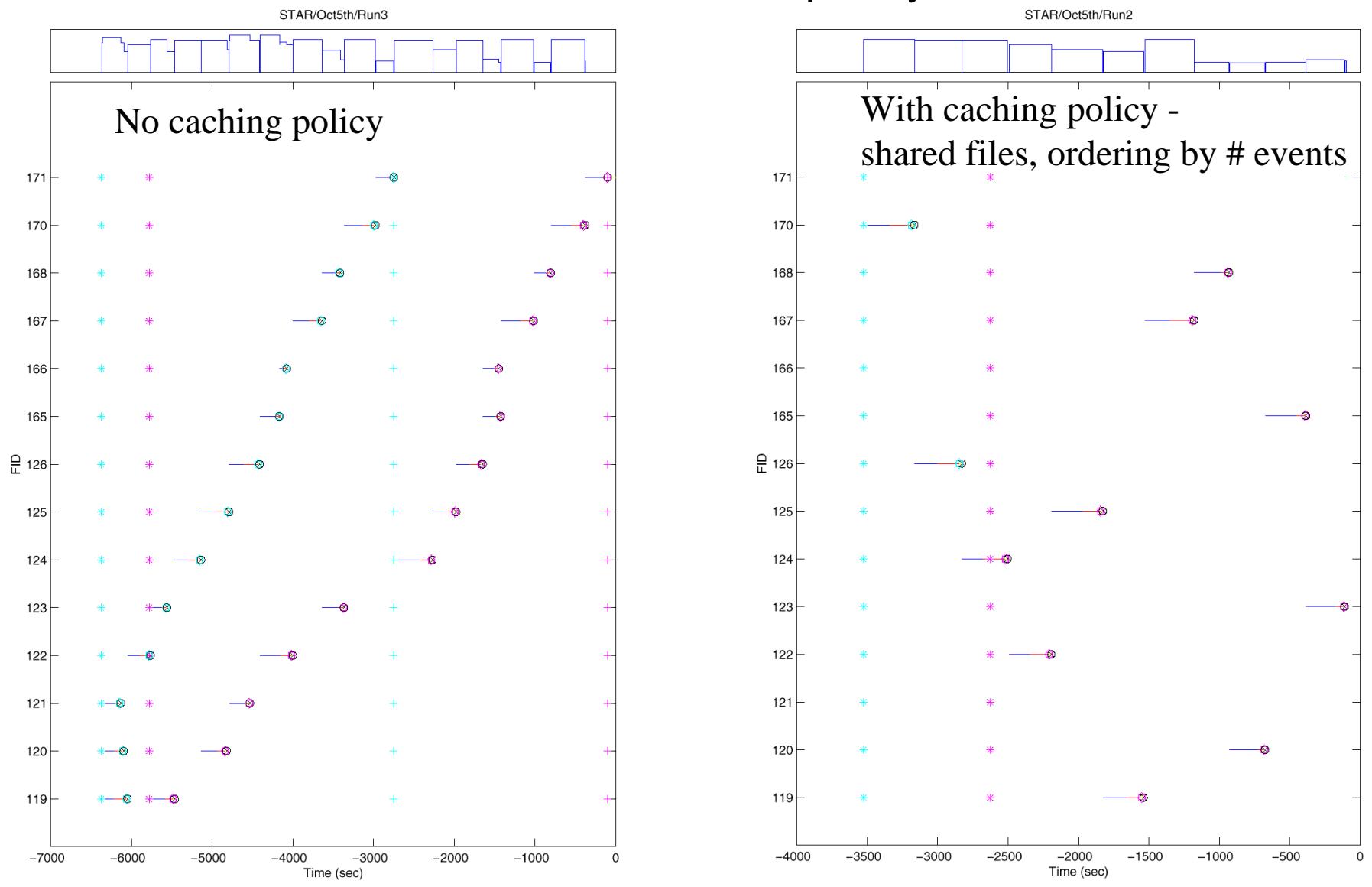
3 queries

3 queries with some shared files, time delay between each query, then the same 3 queries are repeated simultaneously.

The cache was large enough to hold all files so the second time all queries run at processing speed rather than I/O speed.



Shared access policy



Time: 21:43:34 (-8014) Cache: 0 MB FID: 120 File Size: 85

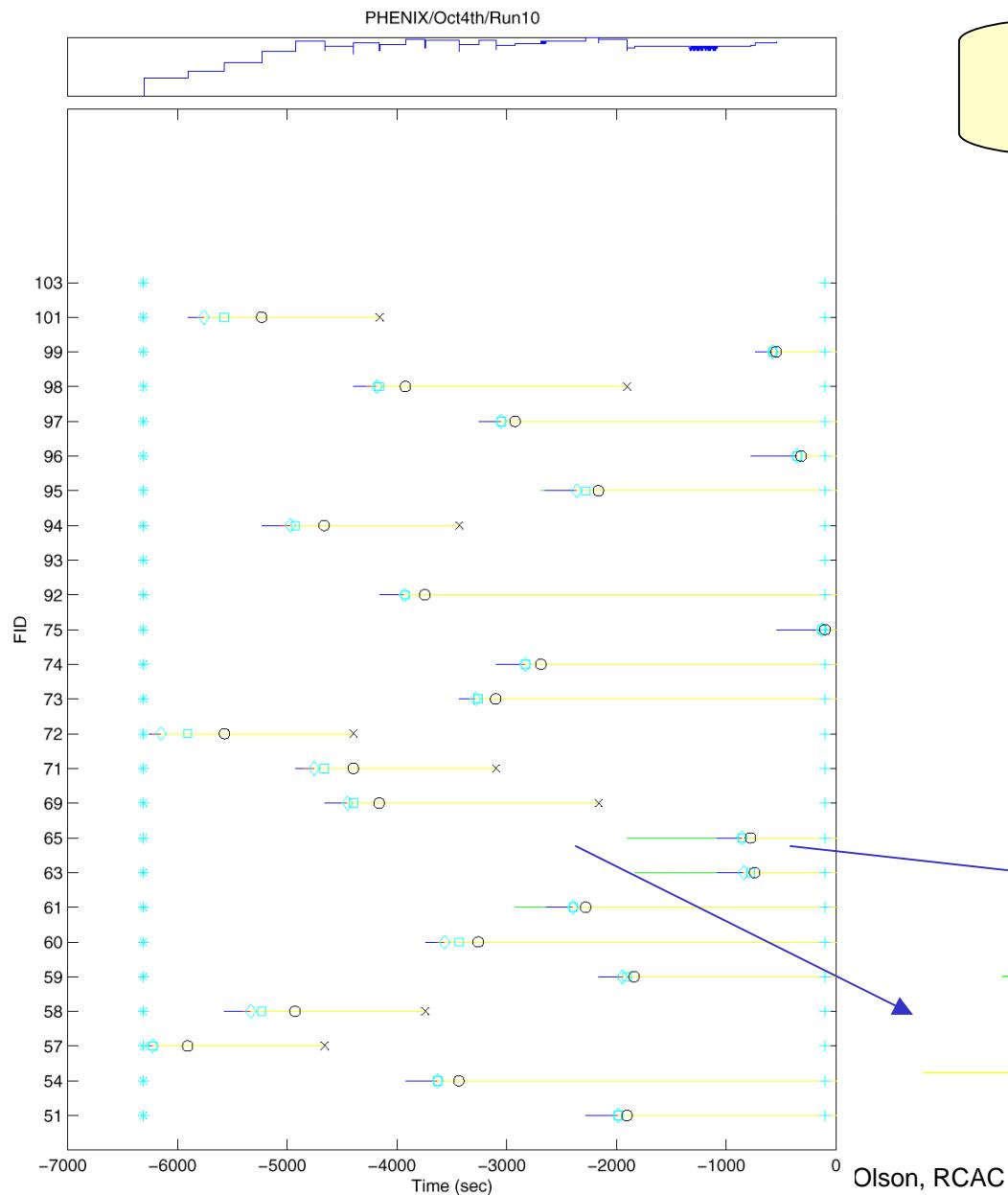
4 Dec 1998

Time: 12:50:36 (-3459) Cache: 302 MB FID: 122 File Size: 255

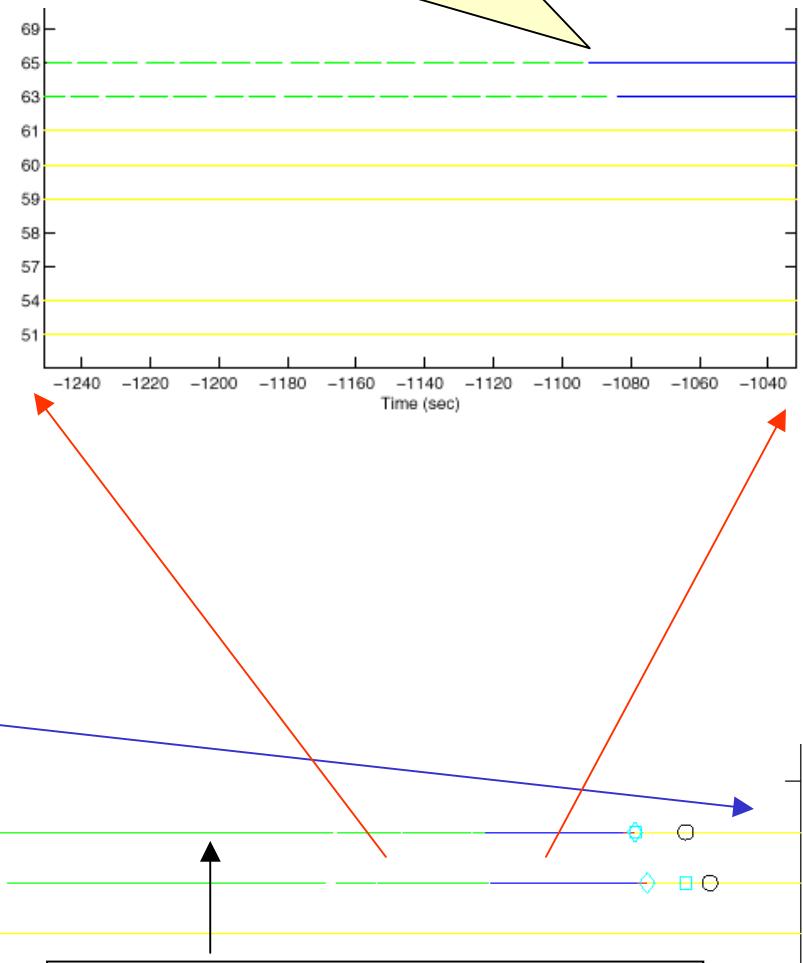
D. Olson, RCAC

13

Detail



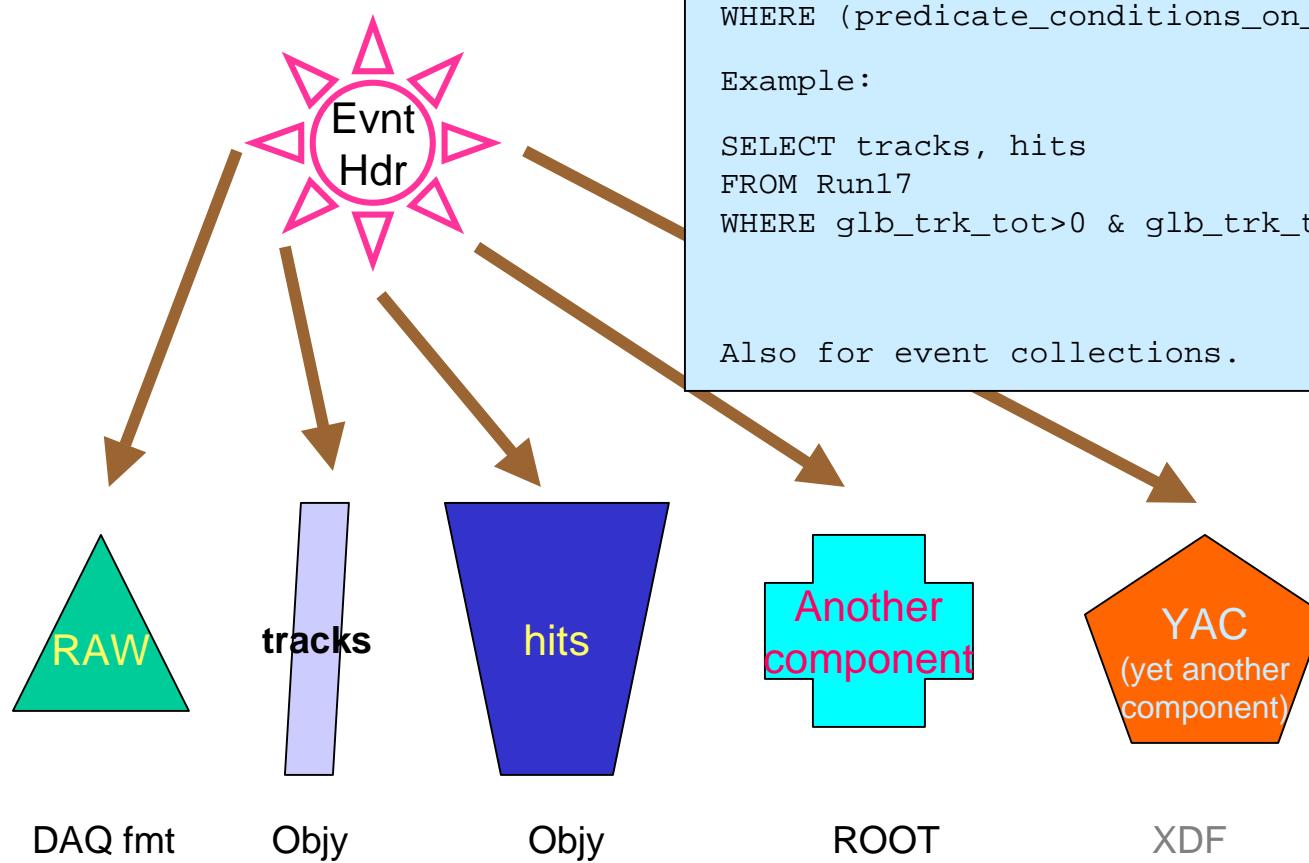
HPSS recovered
& pftp succeeds again



GC software features for MDC2

- Multiple event components
 - introduce “file bundles” - a set of events and their corresponding files
- Files in multiple directories (use Objy catalog)
- Time estimate for query
- parallel query execution (multiple iterators, same token)
- persistent state storage manager (can restart)
- query object can save event collection
- incremental index updates
- Linux analysis codes (Objectivity permitting)
- See more details at
<http://www-rnc.lbl.gov/GC/email/archive/msg00663.html>

Multiple named components



```
SELECT (component_name, component_name, ...)  
FROM dataset_name  
WHERE (predicate_conditions_on_properties)
```

Example:

```
SELECT tracks, hits  
FROM Run17  
WHERE glb_trk_tot>0 & glb_trk_tot<10 & n_vert_total<3
```

Also for event collections.

Use Objy catalog

- In MDC1 Objy catalog used by cache manager
(but some pieces rely on single cache directory)
- Plan to use Objy catalog for all files, even non-Objy files.
- This provides storage manager with file_ID - file path mapping.
- Still plan to have all files in single HPSS directory
(may need to change this).

Time estimate for query

- Query estimator does “estimation”.
- Quick estimate returns (min., max.) numbers for number of events, number of files (for all components requested).
- Full estimate returns exact numbers
 - plus a time estimate which is approximate minimum time under light load

Parallel query execution

- Query object (QO) defines (creates) query with query estimator (QE),
a query token is created.
- QO will return token, or accept token.
- Separate jobs can be run using same query token --> parallelism.

Summary

- MDC1 results
 - Demonstrated functional software integration
GC-Objy-HPSS-STaF-STAR
 - Analyzed results of simple test cases
 - most aspects understood
 - some interesting behaviors noted
- MDC2
 - Multiple event components (across files)
 - Integration with ROOT
 - Linux
- Beyond MDC2 - TBD
 - Production use
 - Further R&D